A distributionally robust approach to active fault detection for linear stochastic dynamic systems ⋆

Ting Xue [*], **Dong Zhao** [**], **Linlin Li** [***], **Yueyang Li** [****], **Maiying Zhong** [†]

[*] *Shanghai Maritime University, Shanghai, China*
[**] *Beihang University, Beijing, China*
[***] *Beijing University of Science and Technology, Beijing, China*
[****] *University of Jinan, Jinan, China*
[†] *Shandong University of Science and Technology, Qingdao, China*

Abstract: This paper focuses on addressing active fault detection issues for linear stochastic dynamic systems via distributionally robust optimization (DRO) technique. By constructing an observer-based residual generator over finite time horizon, residuals in fault-free and faulty situations are characterized by mean-covariance based ambiguity sets, concerning the unknown exact probability distributions for stochastic disturbances. Then the design of an auxiliary signal for active fault detection is formulated as a DRO problem in the sense of minimizing the auxiliary signal energy while guaranteeing satisfactory false alarm rate (FAR) and missed detection rate (MDR). By bridging the FAR and MDR involved chance constraints with norm-bounded sets, a deterministic formulation of the targeting DRO problem is derived without making any distribution assumptions. Analytical solutions to the optimal auxiliary signal and separation hyperplane are obtained. The proposed method can ensure not only the worst-case fault detection accuracy in the probabilistic context but also the robustness against distributional uncertainties of stochastic disturbances.

Keywords: Active fault detection, distributionally robust optimization, stochastic dynamic system

1. INTRODUCTION

Towards improving system safety and reliability, model-based fault detection (FD) has attracted considerable attention over the past few decades and tremendous achievements have been obtained, see e.g., Chen and Patton (1999); Campbell and Nikoukhah (2004); Ding (2021). Among the involved results, one mainstream technology is passive FD, the key idea of which is generating a so-called residual signal by means of methods such as observers, parity relation and parameter estimation, etc., Chen and Patton (1999). In recent years, great interest in active FD has been growing both in research and engineering committees, which provides a powerful tool for multiplicative fault detection and diagnosis. Different from passive FD where continuous monitoring can be performed, in active FD a testing signal named auxiliary signal is designed and injected into the system during a test period, in such a way possible abnormal behaviors of the system can be separated from the normal mode before the end of the test period Campbell and Nikoukhah (2004).

According to the monitored system subject to either deterministic or stochastic disturbances, active FD results can be roughly classified into deterministic methods and stochastic methods. In deterministic active FD, additive disturbances are generally considered to be norm-bounded (or energy-bounded) and the input-output observations are described with closed sets both for fault-free and faulty modes. An auxiliary signal is thus designed towards separating these sets completely. For example, Zhai et al. (2015) presented a set-membership aided active FD method for open-loop systems with energy-bounded disturbances. Blanchini et al. (2017) proposed a duality based approach via convex programming for active fault isolation purpose. Considering closed-loop systems with deterministic disturbances, Ashari et al. (2012) gave a robust active FD method by designing an input signal in the sense of minimizing the linear quadratic regulator control performance while achieving full separation of abnormal system mode with the normal one. Niemann (2006) proposed an active FD setup based on Youla-Jabr-Bongiorno-Kucera (YJBK) parameterization technique and, on this basis, Wang et al. (2017) extended the results to fault-tolerant control for incipient faults. Taking into account fault detection accuracy and control performance indices

⋆ This work was supported in part by National Natural Science Foundation of China under Grants (62103247, 62233012, 62322303, 62273021, 61973135), Chinese Postdoctoral Science Foundation (2021M702022, 2022T150388), Postdoctoral Program for Innovative Talents of Shandong Province of China (SDBX2021010), Project of Shandong Province Higher Educational Youth and Innovation Talent Introduction and Education Program, Shanghai Science and Technology Plan Project (20040501200), and the Beijing Natural Science Foundation under Grant 4232047 (Corresponding author: Maiying Zhong, e-mail: myzhong@buaa.edu.cn)

simultaneously, issues of active FD and control have been discussed successively, see for instance, Jan Siroky et al. (2011); Punčochář et al. (2015); Ashari et al. (2012); Forouzanfar and Khosrowjerdi (2017).

In the framework of stochastic active FD, statistical analysis and probability theory are widely applied for auxiliary signal design. For instance, aiming to minimize the upper bound of the probability of mode selection error, Blackmore and Williams (2006) applied Bayes risk as optimization target and presented a hypothesis selection based fault isolation scheme for systems subjected to Gaussian distributed noises. Very recently, Guo et al. (2023) extended the results to state-constrained stochastic dynamic systems. A unified design framework of an active fault detector and the controller for closed-loop stochastic systems has also be demonstrated in Šimandl and Punčochář (2009), where the probability distribution for noises was assumed to be known exactly. For closed-loop systems with Gaussian distributed noises, Kwang-Ki K. Kim and Braatz (2013) combined the model-based prediction technique with statistical distance measures (including Kullback-Leibler divergence and Mahalanobis distance) and developed an optimal input design scheme for active fault diagnosis. The connection of this method with generalized likelihood ratio is also discussed, which, to some extent, bridges the gap between deterministic and stochastic active FD methods.

Despite the remarkable progress of active FD, it remains worth mentioning that: 1) a large amount of stochastic active FD studies assume the probability distribution for random noises is known exactly, which is usually not true in practical applications. As a side effect of this, 2) the robustness of the designed active FD system to distributional uncertainties of stochastic disturbances is poor. Note the effectiveness of distributionally robust optimization (DRO) in handling distributional uncertainties Parys et al. (2016); Zymler et al. (2013); Ghaoui et al. (2003)), growing research efforts have been made for distributionally robust passive FD in recent years, see e.g., Shang et al. (2021); Xue et al. (2020); Tzortzis and Polycarpou (2021). While active FD for stochastic dynamic systems suffering distributional uncertainties remains an open topic.

Proceedings of the above observations, in this paper we demonstrate a new active FD approach for linear stochastic dynamic systems in the framework of DRO. Without knowing exact probability distributions for disturbances, an optimal auxiliary signal is designed in the sense of minimizing the energy of auxiliary signal with the worst-case false alarm rate (FAR) and missed detection rate (MDR) being ensured not exceeding predefined levels. Simultaneously, by specifying the probability distributions for disturbances with mean-covariance based ambiguity sets, the designed FD system will be robust to distributional uncertainties of disturbances. The paper is organized as follows. Problem formulation is given in Section II. Section III demonstrates the main results of the paper. Conclusion ends the paper in Section IV.

Notations. In this paper, $\text{diag}\{\cdots\}$ denotes the diagonal matrix with elements $\{\cdots\}$. $\lambda_{max}\{A\}$ represents the largest eigenvalue of matrix A. For a sequence vector $\xi(i) \in \mathbb{R}^n$, we denote the column vector comprised of elements $\xi(i)$, $i = 0, 1, \cdots, N$ by $\mathbf{col}\{\,\xi(i)|_{i=0:N}\,\} = [\,\xi^T(0)\ \xi^T(1)\ \cdots\ \xi^T(N)\,]^T$. Let $\mathbb{P}_\zeta$, $\mathbb{E}[\zeta]$ and $\mathbb{V}[\zeta]$ be the probability distribution, the mean and the variance of random variable $\zeta \in \mathbb{R}^m$, respectively. $\xi \sim (\bar{\xi}, \Sigma_\xi)$ means that variable ξ follows probability distribution with mean $\bar{\xi}$ and variance Σ_ξ. $P \succ 0$ means P is positive definite. Given matrices A, B, C, D and an integer $N > 0$, operators $\mathcal{O}_N$ and $\mathcal{F}_N$ are respectively defined as follows

$$\mathcal{O}_N(A, C) = \left[\, C^T\ (CA)^T\ \cdots\ (CA^N)^T \,\right]^T$$

$$\mathcal{F}_N(A, B, C, D) = \begin{bmatrix} D & 0 & \cdots & 0 \\ CB & D & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{N-1}B & \cdots & AB & D \end{bmatrix}$$

2. PROBLEM FORMULATION

Consider the following stochastic linear discrete time-invariant system modeled as

$$\mathcal{G}_i : \begin{cases} x_i(k+1) = A_i x_i(k) + B_i u(k) + E_i d_i(k) \\ y(k) = C_i x_i(k) + D_i u(k) + F_i d_i(k) \end{cases} \tag{1}$$

where $\mathcal{G}_i$ with $i = 0$ denotes the system mode in fault-free case and $i = 1$ the mode in faulty case, $x_i \in \mathbb{R}^n$, $u \in \mathbb{R}^l$ and $y \in \mathbb{R}^m$ are the system state, auxiliary signal and output vectors, respectively, $x_i(0) = x_{i,0}$, $d_i \in \mathbb{R}^q$ is a zero-mean stochastic disturbance vector without knowing exact probability distribution, $A_i, B_i, C_i, D_i, E_i, F_i$ are system matrices of appropriate dimensions and

$$\begin{bmatrix} A_1 & B_1 & E_1 \\ C_1 & D_1 & F_1 \end{bmatrix} = \begin{bmatrix} A_0 & B_0 & E_0 \\ C_0 & D_0 & F_0 \end{bmatrix} + \Delta_f^{\text{ref}}$$

Δ_f^{ref} represents the reference for multiplicative faults under consideration.

Given a gain matrix L stabilizing $(A_0 - LC_0)$, an observer-based residual generator can be constructed as follows

$$\begin{cases} \hat{x}(k+1) = A_{L,0}\hat{x}(k) + B_{L,0}u(k) + Ly(k) \\ r(k) = y(k) - C_0\hat{x}(k) - D_0 u(k) \end{cases} \tag{2}$$

where $r \in \mathbb{R}^m$ is the residual signal, $\hat{x}(0) = \hat{x}_0$ and

$$A_{L,0} = A_0 - LC_0,\ B_{L,0} = B_0 - LD_0$$

Let $\tilde{x}(k) = [x_i^T(k) \quad \hat{x}^T(k)]^T$. It is easy to derive from (1) and (2) that the dynamics of residual is driven by

$$\begin{cases} \tilde{x}(k+1) = \bar{A}_i\tilde{x}(k) + \bar{B}_i u(k) + \bar{E}_i d_i(k) \\ r(k) = \bar{C}_i\tilde{x}(k) + \bar{D}_i u(k) + F_i d_i(k) \end{cases} \tag{3}$$

where $\tilde{x}(0) = \tilde{x}_0 = [x_{i,0}^T\ \ \hat{x}_0^T]^T$ and

$$\bar{A}_i = \begin{bmatrix} A_i & 0 \\ LC_i & A_{L,0} \end{bmatrix},\ \bar{B}_i = \begin{bmatrix} B_i \\ B_{L,0} + LD_i \end{bmatrix},\ \bar{E}_i = \begin{bmatrix} E_i \\ LF_i \end{bmatrix}$$

$$\bar{C}_i = [\, C_i \quad -C_0 \,],\ \bar{D}_i = D_i - D_0.$$

Despite the unknown exact probability distribution for disturbance d_i and the initial state $\tilde{x}_0$, we can without loss of generality assume

$$\mathbb{E}\begin{bmatrix} \tilde{x}_0 \\ d_i(k) \end{bmatrix} = 0,\ \mathbb{E}\left\{ \begin{bmatrix} \tilde{x}_0 \\ d_i(k) \end{bmatrix} \begin{bmatrix} \tilde{x}_0 \\ d_i(j) \end{bmatrix}^T \right\} = \begin{bmatrix} \Sigma_{\tilde{x}_0} & \\ & \delta_{kj}\Sigma_{d,i} \end{bmatrix}$$

where δ_{kj} is the delta function that satisfies $\delta_{kj} = 0$ for $k \neq j$ and $\delta_{kj} = 1$ for $k = j$, $\Sigma_{\tilde{x}_0} \succ 0$ and $\Sigma_{d,i} \succ 0$.

It is seen from (3) that the residual signal r is coupled with auxiliary input u due to multiplicative fault Δ_f^{ref}. This is different from the case with additive faults where the residual signal is decoupled with system input completely. On the other hand, from the viewpoint of classification, an FD issue can be regarded as a binary classification problem, i.e., the fault-free and faulty classes. So a hyperplane $\mathcal{H}(w,b)=\{r|w^T r=b\}$ with parameters w and $b\in\mathbb{R}$ can be determined in residual evaluation stage, such that, by using $J(r)=w^T r(k)$ as an evaluation function and $J_{th}=b$ as the threshold, the occurrence of a fault can be detected by performing the following decision logic

$$\begin{cases} J(r)>J_{th}, & \text{fault alarm} \\ J(r)\leq J_{th}, & \text{no alarm} \end{cases} \tag{4}$$

Thus the design of an active FD system for process (1) lies in the design of auxiliary signal $u(k)$ and hyperplane parameters w and b.

To assess the fault detection performance, we recall the definitions of FAR and MDR (Ding et al. (2019))

Definition 1. Given an FD system with residual generator (2), evaluation function $J(z)$, threshold J_{th} and decision logic (4), we call the conditional probability

$$\mathrm{P}_{\mathrm{FAR}}=\Pr\{J(r)>J_{th}|\Delta_f^{\mathrm{ref}}=0\}$$

FAR with respect to (w.r.t.) uncertainty d_0, $\tilde{x}_0$, and the conditional probability

$$\mathrm{P}_{\mathrm{MDR}}=\Pr\{J(r)\leq J_{th}|\Delta_f^{\mathrm{ref}}\neq 0\}$$

is called the MDR w.r.t. fault mode $\Delta_f^{\mathrm{ref}}\neq 0$.

Note that, due to the unknown precise probability distributions for initial state $\tilde{x}_0$ and stochastic disturbance d_i, it is difficult to compute the exact value of FAR and MDR in the probabilistic context. Since means and covariance matrices of $\tilde{x}_0$ and d_i are known, an alternative way is to measure the upper bounds of FAR and MDR in the worst-case setting. Meanwhile, the injection of an auxiliary signal should not influence the system dynamics seriously. Thus the energy of $u(k)$ over testing period should be as small as possible with satisfactory FAR and MDR. In this context, the design of the active FD system is formulated as follows.

Problem 2. Given system (1) and residual generator (2), find an appropriate auxiliary signal over finite time horizon, i.e., $u(k)$ with $k\in[0,N]$ $(N>0)$, and determine a hyperplane $\mathcal{H}(w,b)=\{r|w^T r=b\}$, such that for predefined upper bounds of FAR and MDR the energy of auxiliary signal is minimized, i.e.,

$$J_{opt}:=\min_{u,\,w,\,b}\sum_{k=0}^{N}u^T(k)u(k) \tag{5}$$

$$s.t.\begin{cases} \tilde{x}(k+1)=\bar{A}_i\tilde{x}(k)+\bar{B}_i u(k)+\bar{E}_i d_i(k) \\ r(k)=\bar{C}_i\tilde{x}(k)+\bar{D}_i u(k)+F_i d_i(k) \\ \mathbb{E}\begin{bmatrix}\tilde{x}_0 \\ d_i(k)\end{bmatrix}=0, \\ \mathbb{E}\left\{\begin{bmatrix}\tilde{x}_0 \\ d_i(k)\end{bmatrix}\begin{bmatrix}\tilde{x}_0 \\ d_i(j)\end{bmatrix}^T\right\}=\begin{bmatrix}\Sigma_{\tilde{x}_0} & \\ & \delta_{kj}\Sigma_{d,i}\end{bmatrix} \\ \sup\limits_{\Delta_f^{\mathrm{ref}}=0}\Pr\left\{w^T r>b\right\}\leq\alpha \\ \sup\limits_{\Delta_f^{\mathrm{ref}}\neq 0}\Pr\left\{w^T r\leq b\right\}\leq\beta \end{cases} \tag{6}$$

where $\alpha,\beta\in(0,1)$ represent the predefined upper bounds of FAR and MDR, respectively.

3. MAIN RESULTS

In this section, we first give a determinstic formulation of the target problem (5)–(6) by handling the constraints in (6) with DRO technique. Then analytical optimal solutions to the auxiliary signal and the hyperplane are derived for decision making purpose.

3.1 A determistic reformulation

According to (3), the residual signal over finite time horizon $[0,N]$ can be rewritten as

$$r_N=\mathcal{H}_{u,i}u_N+\mathcal{H}_{dx,i}d_{i,N} \tag{7}$$

where $r_N=\mathbf{col}\{r(j)|_{j=0:N}\}$, $u_N=\mathbf{col}\{u(j)|_{j=0:N}\}$, $d_{i,N}=\mathbf{col}\{\tilde{x}_0,d_i(j)|_{j=0:N}\}$ and

$$\begin{aligned} \mathcal{H}_{u,i}&=\mathcal{F}_N(\bar{A}_i,\bar{B}_i,\bar{C}_i,\bar{D}_i) \\ \mathcal{H}_{d,i}&=\mathcal{F}_N(\bar{A}_i,\bar{E}_i,\bar{C}_i,F_i) \\ \mathcal{H}_{\tilde{x},i}&=\mathcal{O}_N(\bar{A}_i,\bar{C}_i) \\ \mathcal{H}_{dx,i}&=[\mathcal{H}_{\tilde{x},i}\ \ \mathcal{H}_{d,i}] \end{aligned}$$

In what follows $\mathcal{H}_{dx,i}$ is assumed to be of full row rank.

In residual evaluation stage, the evaluation function and threshold become

$$J(r)=w^T r_N,\ J_{th}=b \tag{8}$$

so as to achieve FD by performing (4), where $w\in\mathbb{R}^{m(N+1)}$ and $b\in\mathbb{R}$ are parameters to be designed.

Let $\bar{d}_{i,N}=\mathbb{E}[d_{i,N}]=0$ and $\Sigma_{d_N,i}=\mathbb{V}[d_{i,N}]=\mathrm{diag}(\Sigma_{\tilde{x}_0},\Sigma_{d,i},\cdots\Sigma_{d,i})$. Define

$$\xi_i=\Sigma_{d_N,i}^{-\frac{1}{2}}(d_{i,N}-\bar{d}_{i,N}) \tag{9}$$

where $\xi_i\in\mathbb{R}^q$ is a random vector following the probability distribution with zero-mean and identity variance matrix, i.e., $\xi_i\sim(0,I)$. Then $d_{i,N}=\Sigma_{d_N,i}^{\frac{1}{2}}\xi_i$ and

$$r_N=\mathcal{H}_{u,i}u_N+\mathcal{H}_{dx,i}\Sigma_{d_N,i}^{\frac{1}{2}}\xi_i \tag{10}$$

Denote the residual in fault-free and faulty cases respectively by

$$r_{N,0}=r_N|_{\Delta_f^{\mathrm{ref}}=0}=\mathcal{H}_{dx,0}\Sigma_{d_N,0}^{\frac{1}{2}}\xi_0 \tag{11}$$

$$r_{N,1}=r_N|_{\Delta_f^{\mathrm{ref}}\neq 0}=\mathcal{H}_{u,1}u_N+\mathcal{H}_{dx,1}\Sigma_{d_N,1}^{\frac{1}{2}}\xi_1 \tag{12}$$

The corresponding means and covariance matrices are obtained as

$$\bar{r}_{N,0}=\mathbb{E}[r_{N,0}]=0,\ \bar{r}_{N,1}=\mathbb{E}[r_{N,1}]=\mathcal{H}_{u,1}u_N \tag{13}$$

$$\Sigma_{r_N,0}=\mathbb{V}[r_{N,0}]=\mathcal{H}_{dx,0}\Sigma_{d_N,0}\mathcal{H}_{dx,0}^T \tag{14}$$

$$\Sigma_{r_N,1}=\mathbb{V}[r_{N,1}]=\mathcal{H}_{dx,1}\Sigma_{d_N,1}\mathcal{H}_{dx,1}^T \tag{15}$$

respectively. Despite the unknown exact probability distributions for residual in fault-free and faulty cases, we can, without loss of generality, construct the following mean-covariance based ambiguity sets to characterize their distributions, i.e.,

$$\mathcal{P}_0=\left\{\mathbb{P}_{r_N}\in\mathcal{V}_{m(N+1)}|\mathbb{E}_{\mathbb{P}_{r_N}}[r_N]=\bar{r}_{N,0},\ \mathbb{V}_{\mathbb{P}_{r_N}}[r_N]=\Sigma_{r_N,0}\right\}$$

$$\mathcal{P}_1=\left\{\mathbb{P}_{r_N}\in\mathcal{V}_{m(N+1)}|\mathbb{E}_{\mathbb{P}_{r_N}}[r_N]=\bar{r}_{N,1},\ \mathbb{V}_{\mathbb{P}_{r_N}}[r_N]=\Sigma_{r_N,1}\right\}$$

where $\mathbb{P}_{r_N}$ is the probability distribution of r_N, $\mathcal{V}_{m(N+1)}$ represents all valid probability distributions in space

$\mathbb{R}^{m(N+1)}$. In this sense, the last two conditions in (6) can be converted into

$$\begin{cases} \sup\limits_{\mathbb{P}_{r_N}\in\mathcal{P}_0} \Pr\left\{w^T r_N > b\right\} \le \alpha \\ \sup\limits_{\mathbb{P}_{r_N}\in\mathcal{P}_1} \Pr\left\{w^T r_N \le b\right\} \le \beta \end{cases} \quad (16)$$

which are named the distribtuionally robust chance-constraints (DCCs).

To handle the DCCs in (16), the following theorem is referred.

Theorem 3. (Ghaoui et al. (2003)) Given a random variable $\zeta \in \mathbb{R}^n$ following probability distribution $\mathbb{P}_\zeta$ with $\mathbb{E}[\zeta] = \bar{\zeta}$ and $\mathbb{V}[\zeta] = \Sigma_\zeta \succ 0$, define $\mathcal{P} = \{\mathbb{P}_\zeta \in \mathcal{V}_n | \mathbb{E}_{\mathbb{P}_\zeta}[\zeta] = \bar{\zeta},\ \mathbb{V}_{\mathbb{P}_\zeta}[\zeta] = \Sigma_\zeta\}$. Then, for a fixed vector $g \in \mathbb{R}^n$ and a constant g_0, if

$$\forall \zeta \in \Omega_\zeta^\epsilon,\ g^T\zeta > g_0$$

holds with

$$\Omega_\zeta^\epsilon = \left\{\zeta | (\zeta - \bar{\zeta})^T \Sigma_\zeta^{-1} (\zeta - \bar{\zeta}) \le \kappa^2(\epsilon)\right\}$$

where $\kappa(\epsilon) = \sqrt{(1-\epsilon)/\epsilon}$, $\epsilon \in (0,1)$, then

$$\sup_{\mathbb{P}_\zeta \in \mathcal{P}} \Pr\left\{g^T\zeta > g_0\right\} \le \epsilon$$

It is worth emphasizing that Theorem 3 establishes a bridge between a deterministic condition and a DCC while without making any distribution assumptions on random variable. On this basis, let

$$\kappa(\alpha) = \sqrt{\frac{1-\alpha}{\alpha}},\ \kappa(\beta) = \sqrt{\frac{1-\beta}{\beta}}$$

Define the following sets

$$\Omega_{r_{N,0}}^\alpha = \left\{r_N | \left(r_N - \bar{r}_{N,0}\right)^T \Sigma_{r_N,0}^{-1} \left(r_N - \bar{r}_{N,0}\right) \le \kappa^2(\alpha)\right\}$$
$$\Omega_{r_{N,1}}^\beta = \left\{r_N | \left(r_N - \bar{r}_{N,1}\right)^T \Sigma_{r_N,1}^{-1} \left(r_N - \bar{r}_{N,1}\right) \le \kappa^2(\beta)\right\}$$

which, together with (11)–(15) and $\mathcal{H}_{u,0} = 0$, can be further rewritten as

$$\Omega_{r_{N,0}}^\alpha = \left\{ r_N \left| \begin{array}{l} r_N = r_{N,0} = \mathcal{H}_{dx,0}\Sigma_{d_N,0}^{\frac{1}{2}}\xi_0 \\ \xi_0 \sim (0, I),\ \|\xi_0\|_2 \le \kappa(\alpha) \end{array} \right. \right\}$$
$$\Omega_{r_{N,1}}^\beta = \left\{ r_N \left| \begin{array}{l} r_N = r_{N,1} = \mathcal{H}_{u,1}u_N + \mathcal{H}_{dx,1}\Sigma_{d_N,1}^{\frac{1}{2}}\xi_1 \\ \xi_1 \sim (0, I),\ \|\xi_1\|_2 \le \kappa(\beta) \end{array} \right. \right\}$$

The DCCs in (16) can then be equally represented as

$$\begin{cases} \forall r_N \in \Omega_{r_{N,0}}^\alpha,\ w^T r_N > b \\ \forall r_N \in \Omega_{r_{N,1}}^\beta,\ w^T r_N \le b. \end{cases} \quad (17)$$

Obviously, equation (17) shows that if we can find a hyperplane $\mathcal{H}(w,b) = \{r_N | w^T r_N = b\}$ to separate the ellipsoids $\Omega_{r_{N,0}}^\alpha$ for fault-free case ($i = 0$) and $\Omega_{r_{N,1}}^\beta$ for faulty case ($i = 1$) with probability one, i.e.,

$$\Omega_{r_{N,0}}^\alpha \cap \in \Omega_{r_{N,1}}^\beta = \emptyset \quad (18)$$

the upper bounds of FAR and MDR for any $\mathbb{P}_{r_N} \in \mathcal{P}_0$ in fault-free case and $\mathbb{P}_{r_N} \in \mathcal{P}_1$ in faulty case can be guaranteed not larger than α and β, respectively, when the decision logic (4) with evaluation function and threshold in (8) is used.

In this context, the probabilistic optimization problem (5)–(6) can be equally converted into the following deterministic optimization problem

$$J_{opt} := \min_{u_N,\, w,\, b} u_N^T u_N \quad (19)$$

$$s.t.\quad \Omega_{r_{N,0}}^\alpha \cap \in \Omega_{r_{N,1}}^\beta = \emptyset \quad (20)$$

Remark 4. It is notable that sets $\Omega_{r_{N,0}}^\alpha$ and $\Omega_{r_{N,1}}^\beta$ can be regarded as collections of random variable samples r_N having the two-norm not greater than $\kappa(\alpha)$ and $\kappa(\beta)$, respectively and such samples are obtained w.r.t. $\xi_0 \sim (0, I)$ and $\xi_1 \sim (0, I)$, respectively. In other words, $\Omega_{r_{N,0}}^\alpha$ and $\Omega_{r_{N,1}}^\beta$ can be regarded as samples sets of norm-bounded residuals at levels α and β, respectively.

3.2 Solution and algorithm

In this subsection we will apply useful results given in Campbell and Nikoukhah (2004) to address the optimization problem (19)–(20).

At first, we consider the residual suffered disturbance ξ_i in (10) is a norm-bounded deterministic signal both for fault-free and faulty cases, i.e.,

$$r_N = \mathcal{H}_{u,i}u_N + \mathcal{H}_{dx,i}\Sigma_{d_N,i}^{\frac{1}{2}}\xi_i,\ \|\xi_i\|_2 \le p_i,\ i = 0, 1 \quad (21)$$

where $p_0 = \kappa(\alpha)$, $p_1 = \kappa(\beta)$. The corresponding residual signals are then norm-bounded and belong to

$$\bar{\Omega}_{r_{N,0}}^\alpha = \left\{ r_N \left| \begin{array}{l} r_N = \kappa(\alpha)\mathcal{H}_{dx,0}\Sigma_{d_N,0}^{\frac{1}{2}}\tilde{\xi}_0 \\ \left\|\tilde{\xi}_0\right\|_2 \le 1 \end{array} \right. \right\}$$
$$\bar{\Omega}_{r_{N,1}}^\beta = \left\{ r_N \left| \begin{array}{l} r_N = \mathcal{H}_{u,1}u_N + \kappa(\beta)\mathcal{H}_{dx,1}\Sigma_{d_N,1}^{\frac{1}{2}}\tilde{\xi}_1 \\ \left\|\tilde{\xi}_1\right\|_2 \le 1 \end{array} \right. \right\}$$

in fault-free and faulty cases, respectively. Then we have

$$\Omega_{r_{N,0}}^\alpha \subseteq \bar{\Omega}_{r_{N,0}}^\alpha,\ \Omega_{r_{N,1}}^\beta \subseteq \bar{\Omega}_{r_{N,1}}^\beta \quad (22)$$

which means if $\bar{\Omega}_{r_{N,0}}^\alpha \cap \bar{\Omega}_{r_{N,1}}^\beta = \emptyset$, then $\Omega_{r_{N,0}}^\alpha \cap \Omega_{r_{N,1}}^\beta = \emptyset$ holds. So in what follows we focus on solving the following optimization problem towards an optimal design of auxiliary signal u_N

$$J_{opt}^d := \min_{u_N} u_N^T u_N \quad (23)$$

$$s.t.\quad \bar{\Omega}_{r_{N,0}}^\alpha \cap \bar{\Omega}_{r_{N,1}}^\beta = \emptyset \quad (24)$$

Supposing the residual signals in fault-free and faulty cases can both be modeled as (21), according to Campbell and Nikoukhah (2004), we then have

$$\mathcal{H}_{u,1}u_N = \left[\ \kappa(\alpha)\mathcal{H}_{dx,0}\Sigma_{d_N,0}^{\frac{1}{2}}\ \ -\kappa(\beta)\mathcal{H}_{dx,1}\Sigma_{d_N,1}^{\frac{1}{2}}\ \right]\begin{bmatrix}\tilde{\xi}_0 \\ \tilde{\xi}_1\end{bmatrix} \quad (25)$$

Let $G = \mathcal{H}_{u,1},\ \tilde{\xi} = \left[\ \tilde{\xi}_0^T\ \ \tilde{\xi}_1^T\ \right]^T$, $H = [H_0\ \ -H_1]$ and

$$H_0 = \kappa(\alpha)\mathcal{H}_{dx,0}\Sigma_{d_N,0}^{\frac{1}{2}} \quad (26)$$

$$H_1 = \kappa(\beta)\mathcal{H}_{dx,1}\Sigma_{d_N,1}^{\frac{1}{2}} \quad (27)$$

Equation (25) delivers

$$Gu_N = H\tilde{\xi} \quad (28)$$

It is remarkable that for the monitored system with $i = 0$ and $i = 1$, observations $u(k)$ and $y(k)$ over time horizon

$[0, N]$ for each case consists with one of the system modes. That means, for the same observations $u(k)$ and $y(k)$ with $k \in [0, N]$, there exists no $\tilde{\xi}_0$ and $\tilde{\xi}_1$ satisfying $||\tilde{\xi}_0||_2 \leq 1$ and $||\tilde{\xi}_1||_2 \leq 1$ simultaneously. In this regard, the optimal design of auxiliary input u_N with respect to solving (23)–(24) can be further reformulated as (Campbell and Nikoukhah (2004))

$$J_{opt}^d := \min_{u_N} u_N^T u_N \tag{29}$$

$$s.t. \quad \max_{\eta \in [0,1]} \phi_\eta(u_N) \geq 1 \tag{30}$$

where

$$\begin{aligned} \phi_\eta(u_N) &= \min_{\tilde{\xi}_0, \tilde{\xi}_1} (\eta \tilde{\xi}_0^T \tilde{\xi}_0 + (1-\eta)\tilde{\xi}_1^T \tilde{\xi}_1) \\ &= \min_{u_N} (u_N^T P_\eta u_N) \end{aligned} \tag{31}$$

with $P_\eta = G^T(HQ_\eta^{-1}H^T)^{-1}G$, $Q_\eta = \text{diag}\{\eta I, (1-\eta)I\}$ and $\eta \in (0, 1)$.

According to Ashari et al. (2012); Zhai et al. (2015), the optimal solution to problem (29)–(30) can be obtained by using the following theorem.

Theorem 5. Suppose that $\forall \eta \in [0, 1]$, $HQ_\eta^{-1}H^T \succ 0$. The optimal solution of u_N to (29)–(30) is obtained as

$$u_N^* = \frac{1}{\sqrt{\lambda^*}} \frac{v^*}{\|v^*\|_2} \tag{32}$$

where λ^* is the largest eigenvalue of P_η w. r. t. η, i.e.,

$$\lambda^* = \max_{\eta \in [0,1]} \lambda_{max}\{P_\eta\} \tag{33}$$

and v^* is the corresponding eigenvector. At the optimum, we have

$$J_{opt}^d = \frac{1}{\lambda^*} \tag{34}$$

Moreover, since

$$[I \quad -\mathcal{H}_{u,i}]^T (\mathcal{H}_{dx,i}\mathcal{H}_{dx,i}^T)^{-1} [I \quad -\mathcal{H}_{u,i}] \succeq 0$$

holds, sets $\bar{\Omega}_{r_N}^\alpha$ and $\bar{\Omega}_{r_N}^\beta$ are bounded and convex for u_N according to Lemma 3.2.6 in Campbell and Nikoukhah (2004). In this context, a unique hyperplane exists that can, with probability one, separate the fault-free and the faulty classes w.r.t. norm-bounded disturbances $\tilde{\xi}_0$ and $\tilde{\xi}_1$. To determine such a separation hyperplane $\mathcal{H}(w, b)$, denote the optimal values of η, $\tilde{\xi}_0$, $\tilde{\xi}_1$ respectively by η^*, $\tilde{\xi}_0^*$, $\tilde{\xi}_1^*$. According to (21) and (26)–(27), it is obtained

$$\tilde{\xi}_0 = H_0^{-1} r_N, \quad \tilde{\xi}_1 = H_1^{-1}(r_N - Gu_N)$$

Let

$$f(\tilde{\xi}_0) = \eta \tilde{\xi}_0^T \tilde{\xi}_0, \; g(\tilde{\xi}_1) = (1-\eta)\, \tilde{\xi}_1^T \tilde{\xi}_1$$

We further have

$$\begin{aligned} f(\tilde{\xi}_0) &= \bar{f}(r_N) = \eta r_N^T H_0^{-T} \tilde{\xi}_0 \\ g(\tilde{\xi}_1) &= \bar{g}(r_N) = (1-\eta)\,(r_N - Gu_N)^T\, H_1^{-T} \tilde{\xi}_1 \end{aligned}$$

At the optimum of $\phi_\eta(u_N)$ in (31) with respect to $\tilde{\xi}_0 = \tilde{\xi}_0^*$, $\tilde{\xi}_1 = \tilde{\xi}_1^*$, $\eta = \eta^*$ and $u_N = u_N^*$, it holds

$$\begin{aligned} \bar{f}(r_N) &= f(\tilde{\xi}_0^*) = \eta^* r_N^T H_0^{-T} \tilde{\xi}_0^* \\ \bar{g}(r_N) &= g(\tilde{\xi}_1^*) = (1-\eta^*)\,(r_N - Gu_N^*)^T\, H_1^{-T} \tilde{\xi}_1^* \end{aligned}$$

By utilizing the results of Lemma 3.2.7 in Campbell and Nikoukhah (2004), we have

Algorithm 1 Proposed method for auxiliary signal design

1: Set acceptable upper bounds of FAR $\alpha \in (0, 1)$ and MDR $\beta \in (0, 1)$, an appropriate initial state $\tilde{x}_0$, and covariance matrices $\Sigma_{\tilde{x}_0}$ and $\Sigma_{d,i}$ for $i = 0,\ 1$.
2: Compute matrices $\mathcal{H}_{dx,i}$ and $\mathcal{H}_{u,i}$ and construct $\Sigma_{d_N,i}$.
3: Solve the optimization problem (33) for η^*, v^* and λ^* via generalized eigenvalue-eigenvector technique. Compute u_N^* according to (32).
4: Compute $\tilde{\xi}^*$ with (37) and then compute r_N^*, w^* and b^* with (36), (38) and (39), respectively.
5: During the testing period, compute residual $r(k)$ with (2) at each step $k \in [0, N]$ and construct r_N. Then compute $J(r) = (w^*)^T r_N$ and perform (4) to detect the occurrence of a fault.

$$\begin{aligned} w^* &= \frac{\bar{f}(r_N)}{\partial r_N}|_{r_N = r_N^*} - \frac{\bar{g}(r_N)}{\partial r_N}|_{r_N = r_N^*} \\ &= \eta^* H_0^{-T} \tilde{\xi}_0^* - (1-\eta^*)\, H_1^{-T} \tilde{\xi}_1^* \\ &= \left[\, H_0^{-T} \;\; -H_1^{-T} \,\right] \begin{bmatrix} \eta^* I & \\ & (1-\eta^*)I \end{bmatrix} \begin{bmatrix} \tilde{\xi}_0^* \\ \tilde{\xi}_1^* \end{bmatrix} \\ &= \left[\, H_0^{-T} \;\; -H_1^{-T} \,\right] Q_{\eta^*} \tilde{\xi}^* \end{aligned} \tag{35}$$

where $\tilde{\xi}^* = [(\tilde{\xi}_0^*)^T \quad (\tilde{\xi}_1^*)^T]^T$, w^* defines the hyperplane tangent to the sets $\bar{\Omega}_{r_N,0}^\alpha$ and $\bar{\Omega}_{r_N,1}^\beta$, r_N^* is the optimal r_N at the tangent surface. Note from (21) that

$$r_N^* = H_0 \tilde{\xi}_0^* = Gu_N^* + H_1 \tilde{\xi}_1^* \tag{36}$$

which means $H\tilde{\xi}^* = Gu_N^*$. Together with the optimal solution of (31), vector $\tilde{\xi}^*$ is then specified by

$$\tilde{\xi}^* = Q_{\eta^*}^{-1} H^T (HQ_{\eta^*}^{-1}H^T)^{-1} Gu_N^* \tag{37}$$

Substituting $\tilde{\xi}^*$ into equation (35) delivers

$$w^* = 2(HQ_{\eta^*}^{-1}H^T)^{-1} Gu_N^* \tag{38}$$

The hyperplane through the tangent point is thus determined as

$$b^* = (w^*)^T r_N^* \tag{39}$$

We thus have the following lemma.

Lemma 6. Given optimal auxiliary input u_N^* in (32) w.r.t. η^* and residual generator (21), if $\bar{\Omega}_{r_N}^\alpha$ and $\bar{\Omega}_{r_N}^\beta$ are bounded and convex sets, a hyperplane completely separating $\bar{\Omega}_{r_N}^\alpha$ and $\bar{\Omega}_{r_N}^\beta$ is determined as

$$\mathcal{H}(w^*, b^*) = \{r_N | (w^*)^T r_N = b^*\} \tag{40}$$

with w^* and b^* being given in (38) and (39), respectively.

Up to now, by solving the deterministic optimization problem (29)–(30), we have found an optimal auxiliary signal u_N^* and a hyperplane $\mathcal{H}(w^*, b^*)$ that can separate the fault-free and faulty modes of system with norm-bounded disturbances at levels α and β with probability one. Remembering the relation (22), we can conclude that u_N^* and $\mathcal{H}(w^*, b^*)$ also solve the optimization problem (19)–(20). That means when the monitored system is subject to stochastic disturbances without knowing exact probability distribution, hyperplane $\mathcal{H}(w^*, b^*)$ can separate the fault-free and faulty modes before the end of the testing period $[0, N]$, with FAR and MDR not larger than α and β, respectively. The proposed auxiliary signal design method for active FD is summarized in **Algorithm 1**. Additionally, since ambiguity sets $\mathcal{P}_0$ and P_1 specify groups of probability distributions that share same means and covariance matrices, the solution of optimization problem (19)–(20) is thus suitable for any probability distributions delivering $\mathbb{P}_{r_N} \in \mathcal{P}_0$ in fault-free case and $\mathbb{P}_{r_N} \in \mathcal{P}_1$ in faulty case. Namely, the developed active FD system is robust against distributional uncertainties of stochastic disturbances.

Remark 7. It is worth noting that this work connects the stochastic active FD and deterministic active FD with the aid of DRO technique. Without precise probability distributions for disturbances, the relationship between upper bounds of FAR and MDR and the auxiliary signal design is investigated in the probabilistic context, . In the pioneer work of Zhai et al. (2015), a set-membership based active FD scheme was presented and such relations also were discussed even though the concerned disturbances were deterministic and the fault detection rate (FDR) index (i.e., FDR=1-MDR) was computed by using an approximation method. On the other hand, since in the proposed method worst-case FAR and MDR are considered, the designed auxiliary signal is to some extent conservative.

4. CONCLUSION

In this paper, a distributionally robust approach has been proposed to address active fault detection problem for linear stochastic dynamic systems. To this end, an observer-based residual generator is first constructed over finite time horizon. For predefined upper bounds of FAR and MDR, the auxiliary signal design has been formulated as an energy minimization problem subject to FAR and MDR involved DCCs. By virtue of the connections between DCCs and norm-bounded residual sample sets, targeting DRO problem has been reformulated as a deterministic optimization problem that can be solved with generalized eigenvalue-eigenvector technique. An analytical solution to the auxiliary signal has been derived and, on this basis, an optimal hyperplane has been determined for decision making aim. In the future, a data-driven realization of the designed active fault detection system remains worth studying. Besides, with control performance being considered, extending such an approach to stochastic closed-loop systems to achieve an integrated design of fault detector and controller is also an interesting topic.

REFERENCES

Ashari, A.E.E., Nikoukhah, R., and Campbell, S.L. (2012). Active robust fault detection in closed-loop systems: Quadratic optimization approach. *IEEE Transactions on Automatic Control*, 57, 2532–2544.

Blackmore, L. and Williams, B. (2006). Finite horizon control design for optimal discrimination between several models. In *45th IEEE Conference on Decision & Control, San Diego, USA, Dec. 13-15.*

Blanchini, F., Casagrande, D., Giordano, G., et al. (2017). Active fault isolation: A duality-based approach via convex programming. *SIAM Journal on Control and Optimization*, 55, 1619–1640.

Campbell, S.L. and Nikoukhah, R. (2004). *Auxiliary signal design for failure detection.* Princeton Univeristy Press.

Chen, J. and Patton, R.J. (1999). *Robust model-based fault diagnosis for dynamic systems.* Springer.

Ding, S.X. (2021). *Advanced methods for fault diagnosis and fault-tolerant control.* Advanced methods for fault diagnosis and fault-tolerant control.

Ding, S.X., Li, L., and Krüger, M. (2019). Application of randomized algorithms to assessment and design of observer-based fault detection systems. *Automatica*, 107, 175–182.

Forouzanfar, M. and Khosrowjerdi, M.J. (2017). A constrained optimization approach to integrated active fault detection and control. *Iranian Journal of Science and Technology - Transactions of Electrical Engineering*, 41, 229–240.

Ghaoui, L.E., Oks, M., and Oustry, F. (2003). Worst-case value-at-risk and robust portfolio optimization: A conic programming approach. *Operations Research*, 51.

Guo, Y., Liu, Q., Wang, Z., et al. (2023). Active fault diagnosis for linear stochastic systems subject to chance constraints. *Automatica*, 156, 111194.

Jan Siroky, M.S., Axehill, D., et al. (2011). An optimization approach to resolve the competing aims of active fault detection and control. In *50th IEEE Conference on Decision and Control and European Control Conference, Orlando, USA, Dec. 12-15.*

Kwang-Ki K. Kim, D.M.R. and Braatz, R.D. (2013). Optimum input design for fault detection and diagnosis: model-based prediction and statistical distance measures. In *European Control Conference, July 17-19, Zürich, Switzerland.*

Niemann, H. (2006). A setup for active fault diagnosis. *IEEE Transactions on Automatic Control*, 51, 1572–1578.

Parys, B.P.V., Kuhn, D., Goulart, P.J., et al. (2016). Distributionally robust control of constrained stochastic systems. *IEEE Transactions on Automatic Control*, 61, 430–442.

Punčochář, I., Široký, J., and Šimandl, M. (2015). Constrained active fault detection and control. *IEEE Transactions on Automatic Control*, 60, 253–258.

Shang, C., Ding, S.X., and Ye, H. (2021). Distributionally robust fault detection design and assessment for dynamical systems. *Automatica*, 125, 109434.

Tzortzis, I. and Polycarpou, M.M. (2021). Active fault diagnosis for a class of nonlinear uncertain systems: A distributionally robust approach. arXiv:2108.05091.

Wang, J., Zhang, J., Qu, B., et al. (2017). Unified architecture of active fault detection and partial active fault-tolerant control for incipient faults. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 47, 1688–1700.

Xue, T., Zhong, M., Li, L., and Ding, S.X. (2020). An optimal data-driven approach to distribution independent fault detection. *IEEE Transactions on Industrial Informatics*, 16, 6826–6836.

Zhai, S., Wang, W., and Ye, H. (2015). Auxiliary signal design for active fault detection based on set-membership. In *IFAC-PapersOnLine*, 28, 452–457.

Zymler, S., Kuhn, D., and Rustem, B. (2013). Distributionally robust joint chance constraints with second-order moment information. *Mathematical Programming*, 137, 167–198.

Šimandl, M. and Punčochář, I. (2009). Active fault detection and control: Unified formulation and optimal design. *Automatica*, 45, 2052–2059.